

学校编码: 10384

分类号_____密级_____

学号: 24520091153009

UDC _____

厦门大学

硕士学位论文

SCX-IMAC 偶联 LC-MS/MS 鉴定磷酸化
蛋白组模型的建立

Global Profiling of Phosphoproteome Using SCX-IMAC
Coupled with LC-MS/MS

吴雅颖

指导教师姓名: 韩家淮 教授

专业名称: 微生物学

论文提交日期: 2012 年 05 月

论文答辩时间: 2012 年 06 月

学位授予日期: 2012 年 月

答辩委员会主席: _____

评 阅 人: _____

2012 年 06 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ） 2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘 要

蛋白质翻译后修饰包括磷酸化、乙酰化、糖基化和泛素化等，其中，磷酸化是最广泛且目前研究最为透彻的一种翻译后修饰。但是，由于磷酸化蛋白的丰度较低，因此对磷酸化肽段进行有效富集是质谱鉴定成功的关键。常用的磷酸化肽段富集方法有免疫沉淀法、TiO₂层析技术、IMAC 和 SIMAC 等。随着磷酸化蛋白质组学研究的日渐成熟，许多研究小组比较了上述磷酸化肽段富集技术的优劣势，发现 IMAC 对磷酸化肽段的富集效果最好。SCX 是最早建立的多维液相分离技术，其原理是根据肽段的带电量不同，逐步提高盐溶液浓度梯度洗脱 SCX 柱得到逐级分离的样品。近年来，不少实验室将 SCX 和 IMAC 两种方法结合起来，在全细胞磷酸化蛋白组的研究中取得了很好的效果。本论文将详细介绍如何通过 SCX 肽段分离，IMAC 磷酸化肽段富集并采用液相色谱联用质谱仪 Triple TOF 5600 鉴定 RAW 264.7 全细胞磷酸化蛋白组。

我们以约 3mg RAW 264.7 全细胞蛋白为起始样品，采用 FSAP 的方法得到胰蛋白酶酶解肽段，然后将这些肽段进行 SCX 处理得到 12 个组分，接着通过 IMAC 富集每个组分的磷酸化肽段，将富集到的磷酸化肽段脱盐处理后进行 LC-MS/MS 分析，最终得到 10331 个非冗余肽段和 4348 个非冗余磷酸化肽段。为了验证系统的稳定性，我们以 5mg 相同细胞的蛋白重复了上述操作，得到的非冗余肽段数目和非冗余磷酸化肽段数目分别为 12217 个和 6706 个，相同蛋白起始量时，后者比文献报道的肽段数目（约 3900 个）多 71.9%。此外，我们分析了各个组分的磷酸化和非磷酸化肽段的分布，比较了丝氨酸、苏氨酸和酪氨酸磷酸化位点分布，并对丝氨酸和苏氨酸磷酸化的肽段进行 motif-X 基序分析。通过磷酸化位点的数据库搜索，我们还发现了约 3500 个未报道的新的磷酸化位点。

SCX-IMAC 可以和 iTRAQ、SILAC 等标记技术兼容，因此，本论文描述的方法除了能用于定性研究蛋白质的磷酸化外，还为定量磷酸化蛋白组学的研究奠定了坚实的基础。

关键词：蛋白组学；磷酸化；质谱

Abstract

Mass spectrometry (MS) is an important tool to investigate protein phosphorylation. The success in profiling the phosphoproteome by MS-based proteomics has been intimately related to the availability of methods that selectively enrich for phosphopeptides. Strong cation exchange chromatography (SCX) and immobilized metal affinity chromatography (IMAC) are two main phosphopeptides enrichment steps. SCX chromatography separates peptides by solution charge, at pH 3-4, phosphopeptides are expected to elute earlier than their non phosphorylated homologs. IMAC takes advantage of phosphate's affinity for Fe^{3+} to enrich phosphopeptides from the previously collected SCX fractions and is commonly used for phosphopeptides enrichment due to its high affinity for adsorption of phosphopeptides.

We have successfully employed the SCX-IMAC strategy to investigate the phosphoproteome from 3mg and 5mg protein of RAW 264.7. After database searching and manual processing, 4348 and 6706 distinct phosphopeptides were identified, yielded 42.1% and 54.9% recoveries of all distinct peptides. The non-specific binding to peptides and the interference of acidic amino acid in peptides are two primary reasons for low proportion of phosphopeptides. Of these phosphopeptides, 87.2% of the phosphorylation sites identified were on serine residues, 12.8% were on threonine residues, and no phosphorylation on tyrosine was detected. These percentages are similar to those reported in the literature (1800:200:1). Motif-X was used to enrich conserved motifs in identified phosphopeptides, and 79 serine phosphorylation motifs and 22 threonine phosphorylation motifs were identified. Of these motifs, 10 motifs can be found to corresponding kinases.

It is worthy to note that the number of distinct phosphopeptides identified in our study is higher than that previously reported (3900) when use 5mg protein, which is probably due to two points.

First, we used Filter-aided sample preparation (FASP) for sample preparation.

In-gel digestion for MS-based proteomics is extremely robust while in-solution digestion minimizes sample handling and is more easily automated. FASP combines these advantages by solubilizing the protein in sodium dodecyl sulfate (SDS), which is exchanged by 8M urea on a standard filtration device.

Second, we used Triple TOFTM 5600 system for peptide identification. Triple TOFTM 5600 is capable of delivering high quality results of complex proteomics samples at high throughput. The high resolution, high mass accuracy and high sensitivity, realized at the fastest MS acquisition speed, leads to the identification of a total of 10331(3mg protein) and 12217(5mg protein) unique peptides of RAW 264.7 lysate. Another key to such excellent result is the average of 34 MS/MS spectra per second achieved in the analysis.

Considering the low proportion (42.1% and 54.8%) of phosphopeptides in total peptides by SCX-IMAC, although considerable number of phosphopeptides has been identified, there are lots of improvements we can make. Such as, (1) Do not incubate for longer than 60 min; (2) Increase the Starting amounts of protein, recommend starting with at least 5 mg of protein; (3) Reduce the amount of IMAC resin.

Key Words: Proteomics; Phosphorylation; Mass Spectrometry

目 录

中文摘要	I
英文摘要	II
中文目录	IV
英文目录	V
第一章 前 言	1
1.1 蛋白质组学	1
1.1.1 蛋白质组学的研究方法.....	1
1.1.2 磷酸化蛋白质组学.....	2
1.1.3 磷酸化蛋白和肽段的富集方法.....	3
1.1.3.1 免疫沉淀富集法.....	4
1.1.3.2 固定化金属离子亲和层析技术.....	5
1.1.3.3 二氧化钛层析技术.....	5
1.1.3.4 SIMAC 技术.....	6
1.1.3.5 磷酸钙沉淀法.....	7
1.1.3.6 化学衍生法.....	7
1.1.4 各种富集方法之间比较.....	7
1.2 分离技术	9
1.2.1 高效液相色谱技术.....	9
1.2.2 反相色谱技术.....	9
1.2.3 亲水相互作用色谱.....	10
1.2.4 静电排斥亲水相互作用色谱.....	10
1.2.5 毛细管电泳分离技术.....	10
1.2.6 多维液相分离技术.....	11
1.3 质谱技术	12

1.3.1	质谱仪的种类.....	12
1.3.1.1	离子阱和四级杆质量分析器.....	14
1.3.1.2	飞行时间质量分析器和与其相关的串联质谱仪.....	14
1.3.1.3	傅立叶变换离子回旋共振和 Orbitrap 质量分析器	15
1.3.2	生物质谱离子化技术.....	15
1.3.2.1	基质辅助激光解析电离.....	16
1.3.2.2	电喷雾电离.....	16
1.3.3	质谱仪的应用.....	17
1.4	AB SCIEX Triple TOF™ 5600 系统	18
1.5	立题背景	20
第二章	材料与方法	21
2.1	实验材料	21
2.1.1	哺乳动物细胞株.....	21
2.1.2	实验相关药品和试剂.....	21
2.1.3	实验室主要仪器.....	22
2.2	实验方法	22
2.2.1	细胞培养.....	22
2.2.2	蛋白样品制备.....	22
2.2.3	蛋白浓度测定.....	23
2.2.4	肽段样品制备.....	24
2.2.5	酶解肽段的 C18 反相色谱柱脱盐.....	25
2.2.6	IMAC 吸附柱制备及磷酸化肽段富集方法	26
2.2.7	SCX 分离磷酸化肽段.....	27
2.2.8	LC-MS/MS 质谱分析	28
2.2.9	质谱数据分析.....	29
2.2.10	磷酸化基序分析.....	29
第三章	结果与分析	30
3.1	研究策略	30

3.2	Bradford 法测定蛋白浓度	31
3.3	FASP 肽段制备法的优势	31
3.4	SCX 色谱图分析	32
3.5	质谱结果与分析	33
3.5.1	SCX 各组分磷酸化肽段和非磷酸化肽段分布比较	34
3.5.2	SCX-IMAC 质谱鉴定结果整合	35
3.5.3	磷酸化位点分布比较	37
3.5.4	多磷酸化肽段分布比较	37
3.5.5	磷酸化基序分析	38
3.5.6	鉴定出的主要磷酸化蛋白及其生物学功能	44
3.6	总结和讨论	45
附录 1	图表索引	47
附录 2	缩略语及中英文对照表	48
参考文献	50
致 谢	56

Table of Contents

Abstract in Chinese.....	I
Abstract in English	II
Contents in Chinese	IV
Contents in English.....	VII
Chapter 1 Introduction.....	1
1.1 Proteomics	1
1.1.1 Research Methods of Proteomics.....	1
1.1.2 Phosphoproteomics	2
1.1.3 Enrichment of PhosphoProteins and phosphoPeptides.....	3
1.1.3.1 ImmunoPrecipitation.....	4
1.1.3.2 Strong Cation Exchange Chromatography	5
1.1.3.3 TiO ₂ Chromatography	5
1.1.3.4 SIMAC Technology	6
1.1.3.5 Calcium Phosphate Secondary.....	7
1.1.3.6 Chemical Derivatization Method.....	7
1.1.4 Comparison of those Enrichment Methods.....	7
1.2 Separation Techniques.....	9
1.2.1 High Performance Liquid Chromatography	9
1.2.2 Reversed-Phase Chromatography	9
1.2.3 Hydrophilic Interaction Liquid Chromatography	10
1.2.4 Electrostatic Repulsion HILIC.....	10
1.2.5 Capillary Electrophoresis Technology	10
1.2.6 Multidimensional Liquid Phase Separation Technology	11
1.3 Mass Spectrometry	12

1.3.1	Types of Mass Spectrometer	12
1.3.1.1	Ion trap and Quadrupole Mass Filter	14
1.3.1.2	Time of Flight Mass Analyzer and related Tandem MS	14
1.3.1.3	FTICR-MS and Orbitrap Mass Analyzer	15
1.3.2	Biological mass spectrometry ionization technology	15
1.3.2.1	Matrix Assistant Laser Desorption Ionization	16
1.3.2.2	Electrospray Ionization	16
1.3.3	Application of Mass Spectrometer	17
1.4	AB SCIEX Triple TOF™ 5600 System	18
1.5	Background of study	20
Chapter 2	Materials and Methods	21
2.1	Experimental Materials	21
2.1.1	Mammalian Cell Lines	21
2.1.2	Reagents	21
2.1.3	Instruments	22
2.2	Material and Methods	22
2.2.1	Cell Culture and Passage	22
2.2.2	Preparation of Protein	22
2.2.3	Determination of Protein Concentration	23
2.2.4	Preparation of Peptides	24
2.2.5	Desaltion of Enzymic Peptides	25
2.2.6	IMAC beads Preparation and Phosphopeptides Enrichment	25
2.2.7	PhosphoPeptides Separation by SCX	27
2.2.8	LC-MS/MS	28
2.2.9	Analysis of MS Data	29
2.2.10	Analysis of Phospho Motif	29
Chapter 3	Results and Analysis	30
3.1	Research Strategies	30

3.2	Determination of Protein Concentration by Bradford.....	31
3.3	Advantages of FASP	31
3.4	Chromatography Analysis of SCX	32
3.5	MS Results and Analysis	33
3.5.1	Comparison of SCX Fraction.....	34
3.5.2	Distribution of Phosphosites	35
3.5.3	Distribution of PhosphoPeptides and non PhosphoPeptides	37
3.5.4	Distribution of multi-PhosphoPeptides.....	37
3.5.5	PhosphoMotif Analysis	38
3.5.6	PhosphoProtein Analysis	44
3.6	Conclusions and Discussion	45
Appendix 1	Index of Figures and Tables	47
Appendix 2	Abbreviation	48
Reference.....		50
Acknowledgements.....		56

第一章 前言

1.1 蛋白质组学

蛋白质组是指由一个基因组、细胞或组织在特定生理或病理状态下表达的所有种类的蛋白质。蛋白质组学（Proteomics，又译作蛋白质体学）的概念最早由 Marc Wilkins¹提出，是系统研究细胞或组织内所有蛋白的鉴定、修饰、定位以及定量的一门学科²。蛋白质组学通过大规模研究蛋白质的特征，以获得蛋白质水平上的关于疾病发生、细胞代谢等过程的整体而全面的认识。2001 年的 Science 杂志把蛋白质组学研究列为六大热点之一，其重要性仅次于干细胞研究，名列第二。然而，蛋白质组学研究是一项复杂且具有挑战性的工作。这是因为，首先，由于转录的可变剪接、蛋白序列的多态性、翻译后修饰（如磷酸化、乙酰化和糖基化等）及其它的蛋白质加工过程，生物体内蛋白质总量远远超出其基因总量；其次，蛋白质在生物体内的浓度范围超越了几乎所有单一分析方法和仪器的动态检测范围。

蛋白质组的研究不仅为生命活动提供了物质基础，也为众多疾病发病机理的阐明及攻克提供了重要的理论根据和解决途径。虽然蛋白质组学的问世时间很短，但其已经在细胞的增殖、分化、异常转化、肿瘤形成等方面进行了有力的探索，涉及白血病、乳腺癌、结肠癌、肾癌、前列腺癌、肺癌、膀胱癌和神经母细胞瘤等，鉴定出许多肿瘤相关蛋白，为肿瘤的早期诊断、药物靶点的发现、疗效判断和预后提供了重要依据。可见，蛋白质组学研究具有十分重要的生物学意义。

1.1.1 蛋白质组学的研究方法

目前，蛋白质组学研究主要有两种策略，即“穷尽法”和“差异法”³。前者采用高通量的研究技术分析生物体内尽可能多的蛋白质，从系统性的角度来看待蛋白质组学，更符合蛋白质组学的本质。后者研究不同时期细胞蛋白组成的变化（如蛋白质在不同环境下的差异表达），以发现差异蛋白种类为主要目标，这种策略更倾向于把蛋白质组学作为研究生命现象的方法和手段。

虽然应用于蛋白质组学研究的技术种类繁多，但是蛋白样品通常只能通过以

下两种方式进行检测：（1）自下而上方式，即进行酶解消化成肽段检测；（2）自上而下方式，即不通过酶解，以原始的完整蛋白形式直接检测⁴。

1.1.2 磷酸化蛋白质组学

蛋白磷酸化是普遍存在于生物体内的一种调节机制，是目前研究最为透彻的一种蛋白翻译后修饰⁵。磷酸化可以改变蛋白质的构象、活性及与其它蛋白之间的相互作用。蛋白质的磷酸化修饰参与了许多生物过程，如转录、翻译调节、代谢、蛋白降解、信号传导、细胞动态平衡、细胞分化、细胞增殖以及细胞存活等⁶。蛋白质的不同位点发生磷酸化可以导致其行使不同的功能。

真核生物中，蛋白磷酸化主要发生在苏氨酸、丝氨酸和酪氨酸残基上⁵。蛋白磷酸化的调控具有严格的时间性和空间性，通常，一个蛋白质有多个磷酸化位点，且不同位点受不同激酶调控。人类基因组的测序结果揭示大约 2-3% 的基因参与编码蛋白激酶⁷。据估计，约 50% 的蛋白质在其生命周期中会发生磷酸化修饰⁵，而人类全蛋白质组中存在大约 100000 个磷酸化位点⁸。蛋白质的磷酸化和去磷酸化是一种普遍的调节机制，是原核生物和真核生物细胞表达调控的关键环节，因此，蛋白质的磷酸化分析和位点鉴定已成为目前蛋白质组学研究的焦点之一⁹。图 1-1 概括了上个世纪 50 年代至 2009 年间磷酸化蛋白质组学研究的重大突破¹⁰。

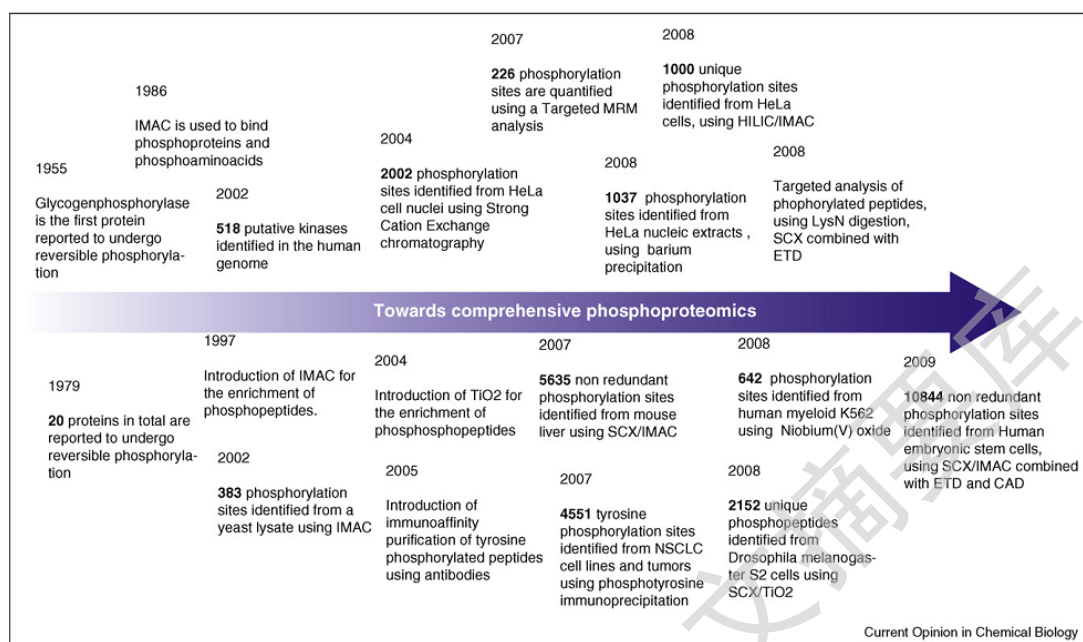


图 1-1 磷酸化蛋白质组学研究的重大突破概况

Fig 1.1 Overview of Important Breakthroughs in Phospho Proteomics

1.1.3 磷酸化蛋白和肽段的富集方法

虽然,得益于质谱技术的飞速发展,蛋白质组学研究的进展十分乐观,但是,由于磷酸化蛋白自身的一些特点,其分析仍极具挑战性。首先,磷酸化蛋白的丰度很低,在特定刺激下仅有少部分蛋白质发生磷酸化;其次,磷酸化的可变性使得蛋白质在不同条件下具有不同的磷酸化形式,且现有的分析方法通常缺乏对磷酸化位点的动态分析,因此部分位点的磷酸化难以鉴定;另外,磷酸酶的存在使样品在制备的过程容易产生去磷酸化现象。因此,对磷酸化蛋白或肽段进行富集以提高其相对含量,是磷酸化蛋白组学研究的重中之重¹¹。如图 1-2¹²所示,根据样品类型和研究目的的不同,目前主要有以下几种磷酸化蛋白和肽段的富集方法。

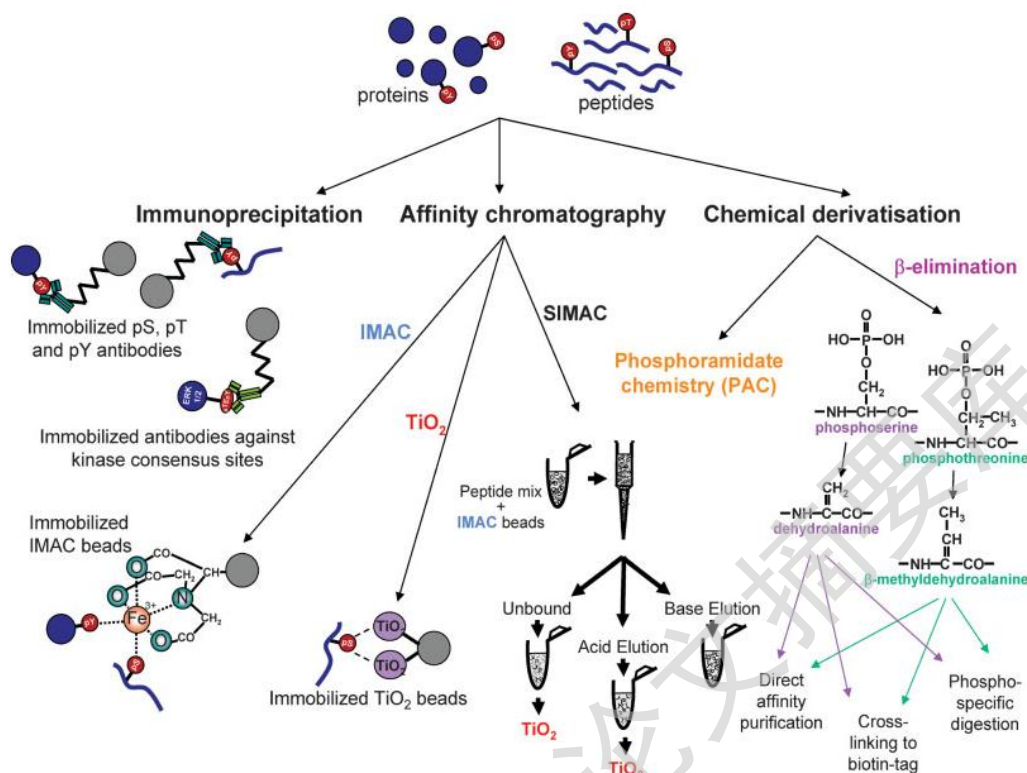


图 1-2 磷酸化蛋白质组学常用富集方法

Fig 1-2 Strategies for phospho specific enrichment

1.1.3.1 免疫沉淀富集法

只关注某一特定的磷酸化蛋白时，通常采用免疫沉淀 (Immunoprecipitation, IP) 的方法将该蛋白从细胞裂解液中纯化出来。经过免疫共沉淀之后的蛋白样品进行 SDS-PAGE 分离后，切下目的蛋白条带，酶解成肽段并脱盐后就可以进行质谱分析鉴定。

据报道，磷酸化丝氨酸、磷酸化苏氨酸、磷酸化酪氨酸在细胞中的比例为 1800: 200: 1¹³，可见，酪氨酸磷酸化的蛋白质在细胞中占的比例非常低，幸运的是，广谱磷酸化酪氨酸抗体的特异性非常高，因此通过免疫共沉淀的方法研究酪氨酸磷酸化的文章层出不穷。而广谱丝氨酸磷酸化抗体和苏氨酸磷酸化抗体的特异性都不高¹⁴，所以研究丝氨酸和苏氨酸的磷酸化时一般不采用此方法。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库